

LIFE EVENTS AND SAMPLE ATTRITION IN THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

Kathleen Short and Edith McArthur, Bureau of the Census

The Survey of Income and Program Participation (SIPP) is a longitudinal survey which began in October 1983. The survey is intended to measure extremely complex phenomena: detailed income sources, reciprocity of Federal and state aid, weekly labor force status, health status and health insurance coverage, taxes, assets, and interest income. In addition to these types of information, the survey collects information on demographic characteristics of all household members. During each visit, interviewers ask questions from a core questionnaire and also, on most visits, ask variable sets of additional questions in the form of "topical modules" on particular issues, such as child care or educational financing.

Approximately 20,000 households formed the "1984 panel." Persons living at the selected addresses were initially interviewed (in four equal size groups) between October 1983 and January 1984. The sample of addresses was selected to be nationally representative of the civilian noninstitutional population of the United States. Persons whose usual residence was at one of the selected addresses were then scheduled for interview once every four months throughout the 2 2/3 years of the 1984 sample's life, that is, into the summer of 1986. In February 1985 and in February 1986, new smaller samples were introduced and a new sample will be introduced each year in the future. More details on both the structure and content of SIPP are available in SIPP Working Paper No. 8401, "An Overview of the SIPP: Update 1" by Nelson, McMillen, and Kasprzyk.

While other major surveys such as the Current Population Survey and the National Crime Survey, both conducted by the Census Bureau, return to the same address for each subsequent visit regardless of whether the occupants of the address change, the SIPP interviewer returns to interview the same persons who form the sample. Persons who move in with SIPP sample persons after the first interviews, while they live with sample persons, are also included in the sample and interviewed. If persons move to a new address, they are followed and interviews are obtained at the new address (for more information about mover's procedures, see Jean and McArthur, 1984.)

Throughout the sample period efforts are made to continue to interview all persons who are ever part of the sample--even if they move to other parts of the country--with a few exceptions: persons who moved into households with sample persons after the first interview are not followed unless they moved with those sample persons; persons who are institutionalized, move outside of the United States, or move to an Armed Forces barracks are not followed; and children under 15 who move and are not accompanied by a sample person who is 15 years old or over are not followed.

Purpose of this Study

Sample maintenance is an important issue in any survey operation. It is especially important when that survey is longitudinal; the Survey of Income and Program Participation (SIPP) follows its sample population through 2 2/3 years. This

paper represents a continuation of the work presented during the 1985 annual meetings of the American Statistical Association (see McArthur and Short, 1985).

The goal of this work is to determine whether attrition from the survey is basically a random phenomenon or whether it is systematic. If attrition is indeed related to personal characteristics, it may be possible to use these results to improve field procedures or adjust weighting specifications. Insofar as improvements are not possible, quantification of individuals leaving the sample is necessary for analysis of succeeding panels of data. If, indeed, attrition is systematic then account should be taken of the degree to which this occurs in statistical analyses of these data. This is so whether the analysis is a simple description of characteristics in a cross-section of the data set involving a subsequent interview or if the analysis is a dynamic one involving several successive interviews at once. Just as individuals with particular characteristics may leave the sample with greater frequency, so individuals experiencing particular events may be more likely to leave the sample. For example, an individual who loses a job may be temporarily absent or may move out of the area, and thus, may not be located by an interviewer for a subsequent wave or waves. If this occurs, any estimation of the probabilities of an event's occurrence over time will be biased to some extent by nonrandom attrition. For a theoretical discussion of this issue see Cox (1959), Gail (1975), Fisher and Kanarek (1974) and Williams and Lagakos (1977).

Thus, this paper describes our findings in several areas: interviewing patterns, reasons recorded for noninterview, and characteristics of sample members by their interviewing status, and finally some initial work studying the relationship between the occurrence of events, such as a change in marital status or employment status, and interviewing status.

Source of Data and Methodology

We define sample attrition as reduction in the numbers of initially interviewed sample persons over the time that those persons are eligible for interview. In a longitudinal survey, disproportionate attrition over time may cause the sample to no longer represent the population from which it was drawn. If, however, the attrition is minimal and no particular subgroups of persons leave the sample, then attrition may not be a cause for worry. We have been exploring the attrition from the SIPP and its potential effect upon the distribution of characteristics of persons remaining in the sample population. The data are the interviewing results from the first five waves of interviewing which covered a twenty-month period from October 1983 through April 1985. The data contained on the extract file that we have used come exclusively from the core portion of the questionnaires--containing information collected during each wave--none of the topical module information is included.

For this work we have included only those sample persons who were 15 years old and over, who were members of a household that lived at one of the selected addresses, and for whom a personal interview was obtained during the first visit by an interviewer. Also we have excluded approximately 16 percent of the original sample who were cut from the sample during a sample reduction conducted as a cost-cutting measure. Further, the sample for this study excludes approximately one quarter of the total SIPP sample for whom the second interview was not scheduled. Because they were not scheduled for all five interviews, they are not included in our current study sample. Persons who became part of the sample after the initial interview, by moving in with initially interviewed persons are also not included in this analysis. The resulting restricted sample consists of 25,138 persons.

Patterns of Attrition

Approximately 79 percent of the sample was interviewed during all five interview waves. About 17 percent were missing at least the fifth interview and may also have missed one or more other interviews, they are called the "leavers." The remaining 4 percent were interviewed during the fifth wave but missed at least one of the intervening interviews, the "returners."

Reasons for Noninterview

Individuals may not have been interviewed for many reasons. An entire household may not have been interviewed, or if the household was interviewed, some individual household members may not have been interviewed. The SIPP is designed to keep a detailed record of the outcome of each interview attempt; detailed codes that represent the outcome of each interview attempt are part of the records kept for each individual and for each household.

Among the reasons recorded for a household noninterview are: no one was home in repeated visits, all household members were away the entire period, the household members refused to be interviewed, the interviewer was unable to locate the unit, roads were impassable, a serious illness or death had occurred in the household, all sample persons in the household were deceased, had moved out of the country, or were living in armed forces barracks, or all sample persons had moved and were living at an unknown address or were living more than 100 miles from a SIPP sampling unit with no available telephone number.

Even if a household interview is obtained, individual household members may have refused to be interviewed or may not have been available during the entire interviewing period and no other household member may have been willing or able to supply information about that individual.

Table 1 displays the reasons for leaving grouped into five categories for the two groups that were not interviewed five times: the "leavers," that is, those who are missing at least the fifth interview; and the "returners," who have a fifth interview but are missing one or more intervening interviews, the returners.

Table 1. Reasons for Noninterviews
(includes persons who became
"out of universe" in the total)

	<u>Leavers</u>	<u>Returners</u>
Total Persons	4222	1038
Percent Those Initially		
Interviewed	16.8	4.1
Reason for leaving	100.0	100.0
<u>Refusals, Total</u>	57.8	38.9
Household	51.3	31.0
Person	6.5	7.9
<u>Person Left Universe</u>	13.4	1.0
Deceased	5.8	-
Institutionalized	3.0	0.4
Armed Forces Barracks	1.6	0.3
Overseas	3.0	0.2
<u>Unable to Contact Household</u>	5.0	19.7
No one home	2.0	6.7
Temporarily absent	2.9	12.9
Unable to locate	-	0.1
<u>Moved Address Unknown</u>	13.4	13.2
<u>Other</u>	10.4	27.2

The principal reason for noninterview in SIPP for both groups was a household refusal. Approximately 13 percent of the leavers actually had "left the SIPP universe," that is, they died, were institutionalized, moved overseas, or moved onto an armed forces barracks and thus they are not true attrition cases. It is useful to point out here that from one wave to the next the reason for a noninterview may change. For example, persons who were not home in one interview may in the next interview have become "refusals." In this paper persons were grouped on the basis of the reason they were not interviewed at the time of the first noninterview.

Relationship between Characteristics & Attrition

In order to determine the effect of attrition from the sample we looked at characteristics of persons by their interviewing experience. Tables 2 and 3 contain these data. We have broken the restricted sample described previously into three subgroups: the "stayers" (those persons who were interviewed during all 5 interviewing waves), the "leavers," and the "returners" (1028 persons not shown separately). Persons who left our universe, that is those who were institutionalized, who died, who moved overseas, etc., are not included in the distributions of leavers or returners. The values of the characteristics shown are as of the first wave. For this analysis we employed chi-square statistics which incorporated a sample design effect. These statistics were calculated to test the independence of the selected characteristics and sample attrition. The calculated chi-square for columns 2 and 3 relevant to a particular characteristic is shown in the table below each characteristic. Those variables determined to have significantly different distributions at the 5 percent level are noted with an asterisk on the table.

Table 2. Distributions of Characteristics of Initially Interviewed Persons by SIPP 1984 Panel

	Restricted Sample	Stayers 5 (interviews)	Leavers (missing at least the 5th interview)		Restricted Sample	Stayers 5 (interviews)	Leavers (missing at least the 5th interview)	
Total Number	25138	19878	3655	Total Number	25138	19878	3655	
WAVE 1 VARIABLES				ETHNICITY:				
REGIONAL OFFICE:				Spanish Origin				5.6
Boston	7.2	7.4	6.9	Not Spanish Origin	94.4	94.7	93.4	
New York	6.9	5.9	10.2	Chi-square columns 2 and 3:	2.36	1 d.f.		
Philadelphia	10.4	10.8	8.2	MARITAL STATUS:				
Detroit	8.4	8.5	8.1	Mar'd, spouse present	58.1	59.9	53.1	
Chicago	7.8	8.4	4.5	Mar'd, spouse absent	0.6	0.5	0.8	
Kansas City	8.4	9.1	5.9	Widowed	7.3	7.4	5.6	
Seattle	8.6	8.9	7.2	Divorced	6.6	6.4	7.3	
Charlotte	8.9	9.2	8.2	Separated	2.3	2.0	3.2	
Atlanta	11.2	10.6	13.5	Never Married	25.0	23.7	30.1	
Dallas	9.7	9.1	12.3	Chi-square columns 2 and 3:	37.21	5 d.f. *		
Denver	5.7	5.8	5.7	HIGHEST GRADE ATTENDED:				
Los Angeles	6.8	6.2	9.4	Less than 9 years	11.4	11.4	8.5	
Chi-square columns 2 and 3:	105.11			9-11 years	16.8	16.5	18.5	
11 degrees of freedom (d.f.) *				12 years	35.8	35.9	36.6	
RESIDENTIAL CHARACTERISTICS				More than 12 years	36.0	36.2	36.4	
Not an SMSA	25.5	26.8	19.3	Chi-square columns 2 and 3:	10.62	3 d.f. *		
SMSA: LT 100,000	1.3	1.3	1.4	EMPLOYMENT RECODE:				
SMSA: 100-249 thou.	9.4	9.6	8.2	With Job:				
SMSA: 250-499 thou.	9.2	9.5	8.0	Worked all weeks	54.6	55.3	54.9	
SMSA: 500-999 thou.	13.4	13.3	13.9	Missed 1+ weeks	1.2	1.2	1.3	
SMSA: 1-2.9 mill.	24.1	23.3	27.9	Time on layoff	0.3	0.2	0.5	
SMSA: 3-14.9 mill.	17.2	16.1	21.5	Job Part of Time:				
Chi-square columns 2 and 3:	52.47	6 d.f. *		No layoff/no looking	1.3	1.3	1.1	
LIVING QUARTERS:				Did look or on layoff	1.3	1.2	1.9	
House, Apt., Flat	94.0	93.9	94.6	No Job:				
Nontransient Hotel	0.2	0.2	0.2	All mo. looked or on layoff	4.6	4.1	6.9	
Perm. in Trans. Hotel	0.1	0.1	0.1	Some looking/layoff	0.6	0.5	0.8	
HU/Rooming House	0.1	0.1	0.1	No looking and no layoff	36.2	36.2	32.7	
Not HU/Room Hse	0.1	0.1	0.2	Chi-square columns 2 and 3:	29.83	7 d.f. *		
Mobile Home, no add.	4.4	4.6	3.8	HOURS WORKED/WEEK:				
Mobile Home, w/add.	0.9	0.9	1.0	Not applicable	37.1	36.6	35.4	
All other	0.3	0.3	-	1 to 19	5.7	5.9	5.0	
Chi-square columns 2 and 3:	10.09	9 d.f.		20 to 34	8.9	8.7	10.0	
LIVING QUARTERS: *				35 to 40	33.3	33.6	34.2	
Owned/Being bought	69.7	71.6	63.5	41 or more	15.0	15.2	15.3	
Rented for cash	28.1	25.9	34.9	Chi-square columns 2 and 3:	3.92	4 d.f.		
Occ'd w/o cash pmt	2.2	2.4	1.6	HOUSEHOLD MONTHLY INCOME:				
Chi-square columns 2 and 3:	41.47	2 d.f. *		Less than 300	4.3	3.9	5.4	
NUMBER OF PERSONS IN HH:				300 to 599	7.7	7.5	7.4	
1 person	11.6	11.5	10.7	600 to 899	8.2	8.1	7.6	
2 persons	29.0	28.8	29.0	900 to 1199	7.8	7.7	8.6	
3 persons	20.3	20.1	21.9	1200 to 1599	11.6	11.3	13.1	
4 persons	20.0	20.8	17.8	1600 to 1999	10.1	10.1	9.5	
5 persons	10.7	10.6	11.1	2000 to 2999	21.7	22.4	21.0	
6 persons	4.4	4.3	4.8	3000 to 3999	13.4	14.0	11.5	
7 persons	2.1	1.9	2.8	4000 or more	15.1	15.0	15.7	
8 or more persons	2.0	2.1	1.9	Chi-square columns 2 and 3:	15.87	8 d.f. *		
Chi-square columns 2 and 3:	11.44	7 d.f.		PERSON MONTHLY INCOME:				
INTERVIEW STATUS:				Less than 300	30.3	29.9	32.7	
Self	67.0	67.9	63.3	300 to 599	16.0	15.8	14.7	
Proxy	33.0	32.1	36.7	600 to 899	12.5	12.3	13.0	
Chi-square columns 2 and 3:	9.97	1 d.f. *		900 to 1199	9.8	9.8	10.5	
INTERVIEW LENGTH: (Minutes)				1200 to 1599	10.3	10.4	10.2	
Less than 15	27.4	26.8	29.3	1600 to 1999	6.7	7.0	5.3	
15 to 29	43.8	44.1	43.7	2000 to 2999	8.7	9.0	8.1	
30 to 44	20.9	21.0	20.1	3000 to 3999	3.1	3.2	3.0	
45 to 59	6.0	6.1	5.3	4000 or more	2.7	2.7	2.7	
60 or more	2.0	2.0	1.6	Chi-square columns 2 and 3:	10.36	8 d.f.		
Chi-square columns 2 and 3:	5.25	4 d.f.		ASSET SUMMARY: Savings account:				
RELATIONSHIP:				Yes	56.5	58.1	50.2	
Reference Person	35.2	35.9	33.1	No	43.5	41.9	49.8	
Primary Individual	13.0	12.7	13.0	Chi-square columns 2 and 3:	25.74	1 d.f. *		
Spouse	28.5	29.7	25.3	All Other Assets:				
Child	16.8	16.3	19.4	Yes	41.1	42.4	36.9	
Other Relative	3.5	3.0	4.8	No	58.9	57.6	63.1	
Non-rel w/rels.	0.4	0.3	0.8	Chi-square columns 2 and 3:	12.69	1 d.f. *		
Non-rel. no rels.	2.5	2.0	3.7	HOUSEHOLD RECEIVES CASH BENEFIT				
Chi-square columns 2 and 3:	41.04	6 d.f. *		Yes	8.2	8.1	7.7	
AGE:				No	91.8	91.9	92.3	
15-24	21.9	21.0	25.6	Chi-square columns 2 and 3:	0.21	1 d.f.		
25-44	37.7	37.9	39.0	HH RECEIVES NONCASH BENEFIT				
45-64	25.2	26.0	24.6	Food stamps	7.2	7.1	6.4	
65 and over	15.1	15.2	10.8	Other only	9.8	9.6	9.8	
Chi-square columns 2 and 3:	24.66	3 d.f. *		No benefits	83.0	83.2	83.9	
SEX:				Chi-square columns 2 and 3:	1.06	2 d.f.		
Male	46.9	46.1	49.0	RACE:				
Female	53.1	53.9	51.0	White	86.8	87.6	84.2	
Chi-square columns 2 and 3:	3.28	1 d.f.		Black	10.4	9.8	12.1	
RACE:				Am. Ind./Esk/Al Nativ	0.4	0.4	0.8	
White	86.8	87.6	84.2	Asian/Pac. Isl.	2.4	2.2	2.8	
Black	10.4	9.8	12.1	Chi-square columns 2 and 3:	13.32	3 d.f. *		
Am. Ind./Esk/Al Nativ	0.4	0.4	0.8					
Asian/Pac. Isl.	2.4	2.2	2.8					
Chi-square columns 2 and 3:	13.32	3 d.f. *						

Table 3. Distributions of Characteristics of Initially Interviewed Persons by Selected Reasons for Noninterview: SIPP 1984 Panel

(Chi-squares separately calculated for stayers (Table 2 column 2) against HH refusal (column 1); against Type D (column 2); against out of Universe (column 3)).

Total Number WAVE 1 VARIABLES	Type D			Type D		
	HH Refusal 2166	(moved address unknown) 564	Out of Universe (died, institu., left country, A.F. barracks) 567	HH Refusal 2166	(moved address unknown) 564	Out of Universe (died, institu., left country, A.F. barracks) 567
REGIONAL OFFICE:						
Boston	7.5	4.3	6.9			
New York	8.9	15.4	6.7			
Philadelphia	7.7	6.9	7.6			
Detroit	9.1	2.7	7.2			
Chicago	4.2	4.6	10.4			
Kansas City	7.6	2.3	7.8			
Seattle	7.8	6.2	7.6			
Charlotte	8.9	6.4	12.0			
Atlanta	12.5	16.1	12.7			
Dallas	12.0	16.5	8.5			
Denver	4.9	8.0	4.8			
Los Angeles	9.0	10.6	7.9			
Chi-square with 11 degrees of freedom against column 1: 48.90* column 2: 78.01* column 3: 7.87						
RESIDENTIAL CHARACTERISTICS						
Not an SMSA	19.5	14.4	24.3			
SMSA: LT 100,000	1.7	-	0.4			
SMSA: 100-249 thou.	8.6	6.0	10.8			
SMSA: 250-499 thou.	8.2	6.7	9.2			
SMSA: 500-999 thou.	14.0	17.2	15.7			
SMSA: 1-2.9 mill.	27.8	31.6	22.2			
SMSA: 3-14.9 mill.	20.1	24.1	17.5			
6 d.f. vs. column 1: 27.90* column 2: 31.62* column 3: 3.50						
LIVING QUARTERS:						
House, Apt., Flat	94.8	94.0	94.0			
All other (collapsed from Table 4 into 1 category)	5.2	6.0	6.0			
9 d.f. vs. column 1: 6.30 column 2: 33.23* column 3: 4.02						
LIVING QUARTERS:						
Owned/Being bought	72.5	26.8	59.4			
Rented for cash	26.3	71.3	38.4			
Occ'd w/o cash pmt	1.2	2.0	2.1			
2 d.f. vs. column 1: 4.62 column 2: 190.66* column 3: 179.8*						
NUMBER OF PERSONS IN HOUSEHOLD:						
1 person	9.8	14.7	23.6			
2 persons	30.6	27.1	36.0			
3 persons	22.8	20.6	14.3			
4 persons	18.1	15.4	11.3			
5 persons	11.3	9.4	7.6			
6 persons	4.1	6.9	3.0			
7 persons	2.4	3.4	2.5			
8 or more persons	0.8	2.5	2.9			
7 d.f. vs. column 1: 12.31 column 2: 9.99 column 3: 36.95*						
INTERVIEW STATUS:						
Self	64.1	68.6	64.7			
Proxy	35.9	31.4	35.3			
1 d.f. vs. column 1: 4.29* column 2: 0.03 column 3: 0.80						
INTERVIEW LENGTH: (Minutes)						
Less than 15	29.0	27.3	30.5			
15 to 29	43.1	44.9	37.9			
30 to 44	20.7	20.6	19.2			
45 to 59	5.5	5.7	8.3			
60 or more	1.7	1.6	4.1			
4 d.f. vs. column 1: 2.14 column 2: 0.27 column 3: 6.54						
RELATIONSHIP:						
Reference Person	35.9	27.0	30.7			
Primary Individual	11.3	20.0	25.7			
Spouse	30.1	13.8	14.5			
Child	17.6	17.9	14.5			
Non-rel w/rels.	0.5	2.7	0.2			
Non-rel. no rels.	1.7	9.4	5.6			
6 d.f. vs. column 1: 2.56 column 2: 118.58* column 3: 74.83*						
AGE:						
15-24	21.1	40.2	21.2			
25-44	36.7	50.2	18.9			
45-64	29.3	8.3	13.9			
65 and over	12.8	1.2	46.0			
3 d.f. vs. column 1: 5.33 column 2: 85.27* column 3: 14.32*						
SEX:						
Male	46.4	54.1	58.4			
Female	53.6	45.9	41.6			
1 d.f. vs. column 1: 0.02 column 2: 4.63* column 3: 11.65*						
RACE:						
White	87.7	75.0	81.0			
Black	9.3	19.9	14.3			
Am.Ind/Esk/AlNativ	0.6	1.2	0.4			
Asian/Pac.Isl.	2.4	3.9	-			
3 d.f. vs. column 1: 1.24 column 2: 27.78* column 3: 8.02*						
ETHNICITY:						
Spanish Origin	3.7	16.8	8.5			
Not Spanish Origin	96.3	83.2	91.5			
1 d.f. vs. column 1: 3.10 column 2: 45.41* column 3: 6.69*						
MARITAL STATUS:						
Mar'd, spouse present	61.8	32.3	39.2			
Mar'd, spouse absent	0.6	1.1	1.9			
Widowed	6.4	2.5	20.8			
Divorced	5.4	13.1	6.9			
Separated	1.8	8.2	3.2			
Never Married	24.1	42.9	28.0			
5 d.f. vs. column 1: 2.82 column 2: 101.90* column 3: 65.83*						
HIGHEST GRADE ATTENDED:						
Less than 9 years	7.5	9.2	19.3			
9-11 years	16.5	27.8	13.9			
12 years	37.8	32.3	29.6			
More than 12 years	38.2	30.7	27.2			
3 d.f. vs. column 1: 10.67* column 2: 16.92* column 3: 53.34*						
EMPLOYMENT RECODE:						
With Job:						
Worked all weeks	56.4	52.8	26.1			
Missed 1+ weeks	1.4	1.4	0.7			
Time on layoff	0.4	0.9	0.2			
Job Part of Time:						
No layoff/no looking	0.9	0.7	0.9			
Did look or on layoff	1.3	3.4	1.2			
No Job:						
All mo.looked or on layoff	5.1	15.6	4.9			
Some looking/layoff	0.4	2.1	0.9			
No looking and no layoff	34.1	23.0	65.1			
7 d.f. vs. column 1: 5.07 column 2: 83.96* column 3: 70.62*						
HOURS WORKED/WEEK:						
Not applicable	35.1	33.5	68.1			
1 to 19	5.4	3.2	3.9			
20 to 34	10.4	10.9	3.9			
35 to 40	34.7	35.1	15.7			
41 or more	14.4	17.4	8.5			
4 d.f. vs. column 1: 3.41 column 2: 4.37 column 3: 77.49*						
HOUSEHOLD MONTHLY INCOME:						
Less than 300	4.8	10.6	8.3			
300 to 599	5.4	14.5	16.9			
600 to 899	7.3	9.4	14.1			
900 to 1199	7.3	12.6	9.5			
1200 to 1599	13.1	13.3	10.6			
1600 to 1999	9.3	9.9	10.6			
2000 to 2999	22.3	18.3	12.7			
3000 to 3999	13.8	3.5	8.5			
4000 or more	16.8	7.8	8.8			
8 d.f. vs. column 1: 9.72 column 2: 60.90* column 3: 56.28*						
PERSON MONTHLY INCOME:						
Less than 300	30.6	40.4	31.9			
300 to 599	13.3	18.8	29.6			
600 to 899	13.4	9.8	15.0			
900 to 1199	10.2	11.2	7.4			
1200 to 1599	10.8	9.4	7.8			
1600 to 1999	5.6	3.2	2.5			
2000 to 2999	9.1	5.1	3.9			
3000 to 3999	3.8	0.7	0.5			
4000 or more	3.0	1.4	1.4			
8 d.f. vs. column 1: 6.65 column 2: 20.97* column 3: 43.02*						
ASSET SUMMARY:						
Savings account:						
Yes	56.5	26.1	48.0			
No	43.4	73.9	52.0			
1 d.f. vs. column 1: 0.61 column 2: 76.18* column 3: 7.93*						
All Other Assets:						
Yes	42.8	14.7	34.6			
No	57.2	85.3	65.4			
1 d.f. vs. column 1: 0.03 column 2: 57.33* column 3: 4.81*						
HOUSEHOLD RECEIVES CASH BENEFIT						
Yes	4.6	17.6	14.3			
No	95.4	82.4	85.7			
1 d.f. vs. column 1: 10.87* column 2: 20.92* column 3: 9.49*						
HOUSEHOLD RECEIVES NONCASH BENEFIT						
Food stamps	2.9	18.6	9.3			
Other only	7.9	14.0	16.0			
No benefits	89.2	67.4	74.6			
2 d.f. vs. column 1: 22.44* column 2: 41.48* column 3: 11.32*						

Among the characteristics for which distributions are shown in the table, we found that regional office, residential characteristics, whether the home was owned or being rented, the sample person's age, race, relationship to reference person, marital status, education, employment status, household monthly income, and asset ownership appear to be related to attrition. Also significant was whether the initial interview was conducted by self or proxy. It is also interesting to note which variables were not significantly different between "leavers" and "stayers;" these are number of persons in the household, sex, ethnicity, length of interview, hours worked per week and reciprocity status (that is, whether the household receives food stamps, WIC and/or AFDC benefits, etc.)

The third table shows distributions of characteristics by selected reasons recorded for leaving. Of particular interest here is the apparent similarity of distributions of characteristics for the household refusals to the "stayers," that is, few of the distributions of characteristics were significantly different. Not surprisingly, distributions of particular characteristics for the persons who moved leaving no followup address and for those who left the universe are quite distinct from the full sample; see for example, the distributions by whether living quarters were owned or being rented, of age of sample persons, of sex, of ethnicity, of marital status, and of person's monthly income. As with Table 2, the calculated chi-squares are shown in the table and those that were significant at the 5 percent level are starred.

Association of Life Events and Attrition

The remaining tables describe relationships between the occurrence of significant life events and continuing in the sample. We hypothesized that the changes we are interested in could be related to the probability of missing interviews.

A number of characteristics which could change during a person's time in sample were examined. These included household size, marital status, reciprocity of cash and noncash benefits, employment status, residence, and household income. For each of these characteristics, with the exception of residence, we compared status as reported during each successive interview to status as reported in the first interview, and change was defined in terms of the interview in which the change was recorded. Then we examined attrition that occurred in the very next interview by whether a change had been recorded in the previous interview. Thus, we examined status as recorded in the second interview and attrition before the third interview; changes recorded in the third interview and attrition before the fourth interview; and changes recorded in the fourth interview and attrition before the fifth interview. We used persons who were interviewed in all five waves as the control group to compare rates of change. Our hypothesis was that those persons who were leaving the sample were more likely to experience a change than those remaining in sample. Measured changes probably underestimate total changes for persons with missing interviews because changes may occur concurrently with leaving the sample, allowing no way of assessing those changes.

Definitions of change used in this analysis and shown in Table 3 and in Table 4 are quite specific in nature and the reader should be aware of the inability of these measures to adequately measure much important change. Household size change was defined as an increase or a decrease in absolute number of household members from one wave to the next. Employment statuses were grouped into three categories: with a job the entire last month of the reference period, with a job part of that month, or with no job during that month. Change in employment status was defined as changing among those three groupings. Household income was grouped into three levels as of the last month of the reference period, and change in total household income was defined as a change of at least \$500 which caused the household income to change level. Change in marital status was defined in terms of any recorded change from the status, such as never married or divorced, as reported in the first interview. Change in reciprocity status of means-tested cash and noncash benefits, such as Aid to Families with Dependent Children and Food Stamps, are also shown in the table. Table 6 shows the relationship between changing residence and attrition. The figures shown are persons who moved during an interview wave and who also became leavers in that wave. Shown for comparison throughout Tables 5 and 6 are the proportions of the "stayers" who experience those events during each interview wave.

For this analysis we used the student's T statistic and incorporated a sample design effect factor. This statistic was used to determine whether changes in the characteristics tested were significantly related to the possibility of attrition before the next interview. An asterisk is shown beside those proportions that are significantly different from the stayer group at the 5 percent level.

Among the characteristics that we examined, change in number of persons in the household, whether an increase or a decrease, appeared to be related to attrition. Also apparently related to attrition were employment status change and change in household income. Not surprisingly, a change in residence was also related to attrition. Interestingly there was no strong relationship between a change in noncash or cash benefit reciprocity and attrition, or between marital status change and attrition. Note that these results are dependent upon the definitions of change implemented in this study. Other definitions may yield different conclusions. However, the testing of more refined definitions awaits future study.

Conclusions and Direction of Future Work

This study suggests that some non-random attrition occurs over the first five waves of the 1984 SIPP panel. Persons leaving the sample have significantly different residential characteristics, age, race, relationship to reference person, marital status, education, employment status, and asset ownership patterns from those who "stay." These results imply that population inferences describing these characteristics using waves beyond the first, without some compensating adjustment to the weights assigned to the remaining sample persons, will be biased to some extent.

Our results further suggest that analyses of probabilities of the occurrence of some events are related to attrition as well. Individuals who experience a household change, job change, change in residence, or change in income are more likely to leave the sample than those who do not. We suggest that probabilities of these particular events will be underestimated if analysis is restricted to persons in the sample through all 5 waves.

Admittedly our measures are crude and fail to take account of many changes that occur and go unobserved by us; changes that occur but do not immediately precipitate withdrawal from sample; and important interactions between personal characteristics, characteristics and event occurrences, and occurrences of several events within the same time period. Further study should, most certainly, incorporate analyses of such interactions.

Meanwhile, all conclusions from analyses of these data should be drawn carefully and made in light of the fact that they are conditional upon an individual experiencing a particular pattern of interview.

REFERENCES

Cox, D.R. (1959), "The Analysis of Exponentially Distributed Lifetimes with Two Types of Failure," Journal of the Royal Statistical Society, Series B, 21, pp. 411-421.

Fisher, L., and Kanarek, P. (1974), "Presenting Censored Survival Data when Censoring and Survival Times may not be Independent," in E. Proschan and R.J. Serfling, Ed., Reliability and Biometry, Society of Industrial and Applied Mathematics, 1984, pp. 303-326.

Gail, M.H. (1975), "A Review and Critique of Some Models in Competing Risk Analysis," Biometrics, 31, pp. 209-222.

Jean, A. and McArthur, E. (1984), "Some Data Collection Issues for Panel Surveys with Application to the Survey of Income and Program Participation," in Proceedings of the Section on Survey Research Methods, American Statistical Association.

McArthur, E., and Short, K. (1985), "Characteristics of Sample Attrition in the Survey of Income and Program Participation," in Proceedings of the Section on Survey Research Methods, American Statistical Association.

Nelson, D., McMillen, D., and Kasprzyk, D. (1985) "An Overview of the Survey of Income and Program Participation: Update 1," SIPP Working Paper Series No. 8401, Washington, D.C.: U.S. Bureau of the Census.

Williams, J.S., and Lagakos, S.W. (1977), "Models for Censored Survival Analysis: Constant and Variable-Sum Models," Biometrika, 64, pp.214-224.

Table 4. Percent of Persons with Changes in Characteristics by Interview of Event Occurrence and Interview Status

Interview Pattern	Total	Change Between Waves			Total	Change Between Waves		
		1 and 2	1 and 3	1 and 4		1 and 2	1 and 3	1 and 4
HOUSEHOLD SIZE								
All Waves	19878	9.1	15.6	20.6				
Only Waves 1,2,3,4	1090			26.1*				
Only Waves 1,2,3	904		19.8*					
Only Waves 1,2	968	12.2						
MARITAL STATUS								
All Waves	19878	1.3	2.7	4.1				
Only Waves 1,2,3,4	1090			5.7				
Only Waves 1,2,3	904		4.6					
Only Waves 1,2	968	2.1						
CASH BENEFITS STATUS								
All Waves	19878	3.3	4.0	4.3				
Only Waves 1,2,3,4	1090			3.3				
Only Waves 1,2,3	904		5.9					
Only Waves 1,2	968	4.1						
NONCASH BENEFITS STATUS								
All Waves	19878	6.3	8.8	9.1				
Only Waves 1,2,3,4	1090			9.8				
Only Waves 1,2,3	904		9.1					
Only Waves 1,2	968	9.1*						
EMPLOYMENT STATUS								
All Waves	19878	10.2	14.1	15.1				
Only Waves 1,2,3,4	1090			17.2				
Only Waves 1,2,3	904		20.4*					
Only Waves 1,2	968	14.6*						
TOTAL HOUSEHOLD INCOME								
All Waves	19878	21.4	24.3	26.6				
Only Waves 1,2,3,4	1090			34.4				
Only Waves 1,2,3	904		31.5*					
Only Waves 1,2	968	31.1*						
WAVE 1 HIGH HHLD INCOME								
All Waves	5761	22.4	23.3	24.4				
Only Waves 1,2,3,4	267			28.8				
Only Waves 1,2,3	232		37.1*					
Only Waves 1,2	221	30.8						
WAVE 1 MID HHLD INCOME								
All Waves	8361	22.8	25.3	28.4				
Only Waves 1,2,3,4	457			39.8*				
Only Waves 1,2,3	368		26.6					
Only Waves 1,2	396	32.8*						
WAVE 1 LOW HHLD INCOME								
All Waves	5555	17.6	22.6	25.0				
Only Waves 1,2,3,4	344			29.1				
Only Waves 1,2,3	294		33.7*					
Only Waves 1,2	327	27.8*						

Table 5. Percent of Persons Who Move in an Interview Wave by Interview Status in that Wave

Interview Pattern	Total	Change Residence during			
		Wave 2	Wave 3	Wave 4	Wave 5
All Waves	19,878	4.8	6.5	6.3	4.8
Only Waves 1,2,3,4	1,090				10.5*
Only Waves 1,2,3	904			11.0*	
Only Waves 1,2	968		7.5		
Only Wave 1	1,058	8.9*			